DOCUMENT RESUME

ED 087 822 TM 003 482

AUTHOR Hanna, Gerald S.
TITLE An Investigation of the "Don't Know" Option in Formative Evaluation.
PUB DATE 74
NOTE 4p.; Paper presented at the Annual Meeting of the National Council of Measurement in Education (Chicago, Illinois, April, 1974)

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *Formative Evaluation; *Guessing (Tests); Multiple Choice Tests; *Response Style (Tests); Summative Evaluation; Test Construction; *Testing Problems; Test Reliability; Test Validity; Test Wiseness

ABSTRACT

Although the "Don't Know" (DK) option has received telling criticism in maximum performance summative tests, its potential use in formative evaluation was considered and judged to be more promising. The pretest of an instructional module was administered with DK options. Examinees were then required to answer each question to which they had responded DK. Concurrent validity of the resulting measures indicated superiority (p. less than .05) of the procedure using the DK option for one of the two criterion measures investigated. Results also revealed a need to teach students to use the option when it is used in formative evaluation. (Author)

An Investigation of the "Don't Know" Option in Formative Evaluation

Gerald S. Hanna
Kansas State University

Abstract

Although the "Don't Know" (DK) option has received telling criticism in maximum performance summative tests, its potential use in formative evaluation was considered and judged to be more promising. The pretest of an instructional module was administered with DK options. Examinees were then required to answer each question to which they had responded DK. Concurrent validity of the resulting measures indicated superiority ($p. < .05$) of the procedure using the DK option for one of the two criterion measures investigated. Results also revealed a need to teach students to use the option when it is used in formative evaluation.

An Investigation of the "Don't Know" Option in Formative Evaluation*

Gerald S. Hanna
Kansas State University

In an attempt to improve reliability by decreasing amount of guessing, a number of standardized tests have contained "I Don't Know" (DK) options in summative multiple-choice tests of maximum performance. There is evidence (e.g., anonymous, 1964; Hanna, 1970) that the DK option fulfills this purpose when tests are scored for number right. However, scoring for number right is objectionable when students have been told to mark a DK option if they don't know an answer; it violates the ethical issue "that no examinee should ever be able to increase the most probable value of his test score by disobeying the examiner's instructions" (Cureton, 1966, p. 44) and seems contrary to the spirit of the relevant section (E. 2.51) of Standards for Educational and Psychological Tests and Manuals, (French, Michael, et. al., 1966, p. 33). Fundamentally, the DK option is unsuitable in tests of maximum performance because it can be faked. Where motivation exists to obtain high scores, test wiseness and personality variables relevant to propensity to guess will significantly influence scores; hence, construct validity is attenuated (Hanna, 1970; Mehrens and Lehmann, 1973, p. 285).

However, Mehrens and Lehmann soundly suggested that when tests are used in formative evaluation to assist in instructional guidance of students, and not for grading purposes, then students may be motivated to be candid in admitting ignorance and marking DK rather than guessing wildly. "To guess correctly could result in a student's being placed in an instructional sequence at too high a level " (1973, p. 286).

Objective

Without in any way defending, supporting or encouraging the use of DK options in summative evaluation, the present study was designed to study empirically the value of the DK option in a setting of formative evaluation as encouraged by Mehrens and Lehmann.

Procedures

The introduction to an instructional module designed to enable students to discriminate between empirical statements and value assumptions was presented to the thirty-seven members of an undergraduate educational psychology class during the summer of 1973. After informing the students

*Paper read at the Annual Meeting of the National Council on Measurement in Education, Chicago, Illinois, April, 1974.

that the module's objective would be assessed on a regular class quiz, the instructions stated:

> The purpose of this pretest is to see if you already have mastered this objective. If you have, there is no need to waste time in studying the module. If you have not mastered the objective, you will want to know this so you can work through the module before Quiz 1.

Examinees then took the ten-item pretest under the following directions:

> For each statement, underline the words to indicate that the sentence is an example of an empirical statement, an example of a value assumption, or that you don't know which it is.

Immediately after individually turning in this pretest, students who answered any question DK, were given another copy of the pretest and these instructions:

> Below is another copy of the pretest. Your instructor has just circled the numbers of the questions you answered "don't know." Please answer each of these questions either "empirical statement" or "value assumptions." You must respond to every question.

Next, subjects were given a criterion test that contained six alternate choice items paralleling the content and format of the pretest and six multiple-choice items, some of which were quite similar in content to those in the pretest (e.g., "Which of the following is a value assumption?") and some of which were less similar in content (e.g., dealing with the possible truth or falsity of empirical statements). None of the criterion items contained a DK option.

Results and Conclusions

Table 1 shows the findings. Perhaps the most salient feature is the very small mean DK response frequency on the pretest. The procedures were clearly unsuccessful in stimulating a significant utilization of the DK option. Consequently, any expected superiority of Variable 1 over Variable 3 in predicting criterion scores would likely be attenuated. As Mehrens and Lehmann (1973, p. 286) observed, "Of course pupils have to be taught that it is to their advantage to use this option rather than guess blindly at an answer." Perhaps greater explicit emphasis of the formative nature of the test would have aided examinees in overcoming the tendency to avoid the DK option--an adaptive behavioral pattern in summative tests.

The greater correlation of Variable 1 than Variable 3 with the more similar criterion (Variable 4) ($p < .05$ for one-tailed Hotelling test for correlated coefficients of correlation) supports the notion that the DK option facilitates formative evaluation. The absence of a superiority of

Variable 1 over Variable 3 in predicting the less similar criterion (Variable 5) fails to support the value of the DK option.

Table 1

Means, Standard Deviations, and Intercorrelations (N=37)

Variables	M's	SD's	r's 2	3	4	5
1. Initial Pretest Raw Score	7.0	1.4	-.33	.98	.48	.37
2. DK Response	0.2	0.4		-.18	-.29	-.20
3. Final Pretest Raw Score	7.1	1.2			.43	.38
4. Alternate-Choice Criterion	5.0	1.5				.42
5. Multiple-Choice Criterion	3.5	1.8				

Collectively, these findings can be interpreted to lend sufficient support to the use of DK options in formative evaluation to merit further study and to suggest that more attention be devoted to teaching students to use the option when it is appropriate.

The value of this study may consist of (1) providing an empirical test of the value of the suggested use of DK options in formative evaluation, (2) providing mildly encouraging, but certainly not compelling, evidence favoring the option's use, and (3) inadvertently demonstrating that considerable emphasis may be necessary before students who are testwise in summative evaluation will be induced to utilized DK options in formative evaluation.

References

Anonymous. E&A "DK" Study. New York: Test Department, Harcourt, Brace, Jovanavich, Inc., 1964.

Cureton, E. F. The Correction for Guessing. The Journal of Experimental Education, 1966, 34, 44-47.

French, J. W., Michael, W. B., et. al. Standards for Educational and Psychological Tests and Manuals. Washington D. C.: American Psychological Association, 1966.

Hanna, G. S. Influence of Directions Regarding Guessing on Validity. Measurement and Evaluation in Guidance, 1970, 3, 35-40.

Mehrens, W. A. and Lehmann, I. J. Measurement and Evaluation in Education and Psychology. New York: Holt, Rinehart and Winston, Inc., 1973.